

HP Operations Orchestration

For the Windows® and Linux® operating systems

Software Version: 9.05

Installing Operations Orchestration with Oracle Real Application Clusters

Document Release Date: June 2012

Software Release Date: June 2012



Legal Notices

Warranty

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notice

© Copyright 2012 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe™ is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Documentation Updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent updates or to verify that you are using the most recent edition of a document, go to:

<http://h20230.www2.hp.com/selfsolve/manuals>

This site requires that you register for an HP Passport and sign in. To register for an HP Passport ID, go to:

<http://h20229.www2.hp.com/passport-registration.html>

Or click the **New users - please register** link on the HP Passport login page.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HP sales representative for details.

Support

Visit the HP Software Support Online web site at:

<http://www.hp.com/go/hpsoftwaresupport>

This web site provides contact information and details about the products, services, and support that HP Software offers.

HP Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Manage support contracts
- Look up HP support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HP Passport user and sign in. Many also require a support contract. To register for an HP Passport ID, go to:

<http://h20229.www2.hp.com/passport-registration.html>

To find more information about access levels, go to:

http://h20230.www2.hp.com/new_access_levels.jsp

Contents

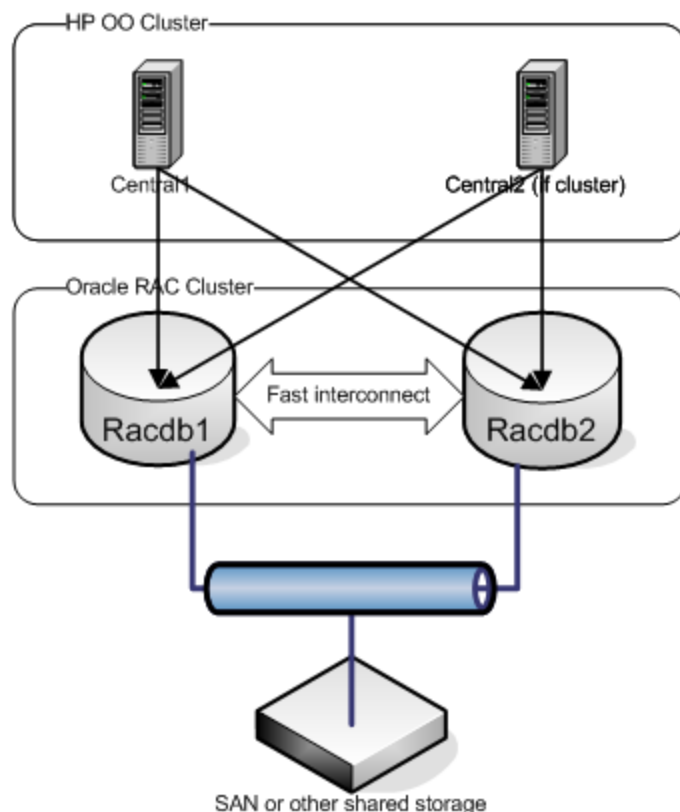
Installing Operations Orchestration with Oracle Real Application Clusters ...	1
Contents	5
Overview	6
HP OO Configuration	7
Linux	7
Windows	8
Connection Failover	9
Performance	10
HP OO Scheduler	11
Troubleshooting	12
JDBC Connection Test	12
RAC Database	12
ONS	12
Connection Load Balancing	13
Useful Information	14

Overview

Oracle Real Application Clusters (RAC) technology is a high-availability solution allowing multiple Oracle database nodes to share the workload and to perform automatic failover in case of failures. It is a hot-hot setup, relying on shared storage and a fast private interconnect to keep all the nodes in the cluster up to date.

Note that Oracle itself does not manage the high availability of the shared storage; the storage must be made highly available through other means, such as SAN or NAS. Also, the RAC technology is not suitable for disaster recovery, as all nodes in the cluster are in proximity of each other. Other technologies such as Oracle DataGuard can be used to enable a hot-standby configuration with databases located away from each other.

The following diagram shows an example of an RAC configuration:



HP OO can be configured to use an Oracle cluster as its database and perform fast connection failovers. HP OO uses the Oracle JDBC driver as a load-balanced connection pool, and handles database node failures by retrying failed transactions through fresh connections. This permits flows to run to completion with minimal time delays.

HP OO Configuration

There are some extra configuration steps that are required for HP OO in order to recognize the database as an Oracle cluster and enable fast connection failover.

An important configuration difference from regular Oracle installations, is the fact that the host to connect to should be specified as a virtual IP address, as opposed to a regular host name. Virtual IP addresses are managed by the cluster software; in case of a node failure, the virtual IP address of a node is automatically transferred to another node. Thus, only virtual IP addresses should be used when configuring the database host.

A second important item particular to clustered environments is configuring HP OO to subscribe to Oracle Notification Service (ONS) events. ONS is a service running on the cluster which produces events known as Fast Application Notification (FAN) events. These events are produced whenever there are changes in the service status on any of the nodes in the cluster (for instance a node going down, or coming back up). ONS also publishes load-balancing advisory events, so that receivers can take advantage of them in order to redirect requests to the best suited services in the cluster. The advantage of these events is that clients do not need to wait for regular TCP timeout events (which could take minutes) before they can conclude about the status of a cluster node.

Linux

The Linux distribution is configured through the **install.config** file. For regular (one-node) installations, the **DB_HOST** and **DB_PORT** identify the host and port of the single instance Oracle. For RAC installations, these settings are not used and will be ignored if populated. Instead, the settings for **ORACLE_RAC_NODES** and **ORACLE_RAC_ONS_CONFIG** must be used. For example:

```
ORACLE_RAC_NODES=racdb1-vip:1521,racdb2-vip:1521
```

```
ORACLE_RAC_ONS_CONFIG=racdb1-vip:6200,racdb2-vip:6200
```

Where **racdb1-vip** is the Virtual IP of the first node in the cluster, while **racdb2-vip** is the Virtual IP address of the second node in the cluster; add more if the configuration involves more than two nodes. The default Oracle (11g Release 2) ports are 1521 for the instances and 6200 for ONS. The ONS ports can be found on each node in the file **\$ORA_CRS_HOME/opmn/conf/ons.config** under the setting **remoteport**.

Note: The **DB_NAME** setting must be used to name the Oracle Service Name for the cluster database. For example: RACDB.

Based upon these settings, the installer will build a JDBC URL that looks like this:

```
jdbc:oracle:thin:@\
```

```
(DESCRIPTION=\  
(LOAD_BALANCE=on) \  
(ADDRESS_LIST=\  
(ADDRESS=(PROTOCOL=TCP) (HOST=racdb1-vip) (PORT=1521)) \  
(ADDRESS=(PROTOCOL=TCP) (HOST=racdb2-vip) (PORT=1521))) \  
(CONNECT_DATA=(SERVICE_NAME=RACDB)) )
```

This type of connect string tells the JDBC driver to use load balancing between the nodes in the address list and to use the RACDB service.

Windows

The Windows installer does not support installing OO with RAC. It is possible, however, to target the installer at one of the nodes, then build a connect string like the one in the previous section and replace it wherever such connect strings are needed. Installing the product on only one node will work fine, because the database is shared, so the tables will be visible to both nodes. The hard part is identifying all the places where the JDBC URL is being used and replacing the current value with the big connect string. These files are:

```
%ICONCLUDE_HOME%\Central\WEB-INF\classes\hibernate.properties
```

```
%ICONCLUDE_HOME%\Central\conf\Central.properties
```

Connection Failover

When HP OO is configured to work with an RAC, its JDBC connection pool listens for FAN events and cleans up those connections that were made to the failed node or instance. All transactions are in progress at the time the node or instance fails, incur exceptions.

HP OO recognizes the error code sent by the JDBC connection pool, and retries the failed transaction up to 10 times, with a time lag between retries specified by the configuration variable called **dharm.hibernate.persistor.connectionRetryTimeSeconds**, which defaults to 10s. If the failover is expected to take more than 100s, the value of this parameter can be increased to accommodate the required failover time. (Oracle CRS requires about 60 tries with 1 second space in between to detect that a node is clearly down.)

Note that HP OO retries transactions whenever it detects any kind of communication failure on the JDBC connection channels. This includes the case when the network connection fails (which is different from the case when a node fails). In this particular case, depending on when the connection failure occurred, the retried transactions may not succeed. For example, a commit is initiated and makes it to the database node, but the reply does not make it back. In that case, there is a high likelihood of failure of a repeated transaction, due to database integrity constraint violations or versioning conflicts. Another case of unrecoverable failure is when a transaction on a node fails, but another user connected to a good node makes an update that becomes incompatible with the replayed transaction. In this case, there is nothing that HP OO can do, as the same situation can occur with a single node configuration.

Performance

Using an RAC has the benefit of scalability through connection load balancing. However, complications arise due to the fact that the nodes in the cluster share a common database (set of files) and communicate heavily through the private interconnect.

Each instance (set of Oracle processes) in a cluster can read and write from the common database. Each instance has a separate SGA (System Global Area), which contains the in-memory data structures needed to run the system as if it were an individual system. Unlike an individual setup however, a technology called Cache Fusion permits the buffer caches from each node to “fuse” and form a global, distributed buffer cache; when a cluster node requires a block, it looks it up in this global cache. Blocks that are not available on the requesting node need to be transferred from the node that has them over the private interconnect. In addition to the traffic generated by the traveling blocks, complex synchronization and locking occurs.

Thus, the performance is not simply a measure of the performance of the shared cluster file system and private interconnect; there are also other subtle interactions. For example, a seemingly trivial task such as generating sequence numbers becomes complicated, as nodes compete to acquire locks against the generator.

HP OO Scheduler

The HP OO Scheduler runs on top of the Quartz third-party application (<http://www.opensymphony.com/quartz/>). The resolution for the bug identified at <http://jira.opensymphony.com/browse/QUARTZ-341> recommends that the Scheduler not be load balanced over many nodes.

Thus, the Scheduler needs to be configured to use only one RAC Virtual IP (e.g., racdb1-vip). Since the Virtual IP gets transferred to an available node upon failure, the scheduler is still highly available.

Another change from the previous scheduler implementation for Oracle, is the fact that a connection pool is being used, whereas the past versions would create a fresh connection every time one was needed. This increases the efficiency of the database access; however the pool size needs to be configured properly (edit the file **\$SCHEDULER_HOME/Scheduler/WEB-INF/jdbcDatasource.xml** and change the **MaxLimit** property).

The Quartz scheduler is not currently coded to perform automatic failover when a node fails in the middle of a transaction, like Central does. Therefore, upon a node failure, some schedules may fail. However, the Scheduler should continue to work as the VIP is transferred to the live nodes, automatically, for the next transactions.

Troubleshooting

JDBC Connection Test

The tool **TestOracleConnection.jar** can be used to test connection to each of the nodes in a cluster, plus the load-balanced connection. For instance, specifying the parameter nodes as follows: `nodes=racdb1-vip:1521,racdb2-vip:1521`, makes the tool test individual connections to `racdb1-vip`, `racdb2-vip` and a connection string consisting of both hosts.

This tool can also be used to subscribe to ONS events by invoking it as follows:

```
$ java -Doracle.ons.debug=true -jar TestOracleConnection.jar\  
subscribe_ons nodes="racdb1-vip:6200,racdb2-vip:6200"
```

RAC Database

On each of the RAC nodes, the following commands can be executed, as the oracle user, to verify the health of the system:

```
srvctl status database -d REDVMRAC
```

The above reports the status of the instances of each node in the cluster.

```
crs_stat [-t]
```

The above reports the status of all the cluster services: ONS, VIP, GSD, listener, instance, database, etc. The `-t` option prints everything in a tabular, more compact view.

ONS

On each of the RAC nodes, verify that ONS is running by issuing the command `onsctl ping` (as Oracle user), and if it reports that it is not running, start it with `onsctl start`.

Beware of the following: if `onsctl ping` reports that **ons** is not running, but the `crs_stat` command reports that the **ons** service is online, you may have some environment variables crossed. Namely, if `onsctl` is seen from `$ORACLE_HOME/bin` instead of `$ORA_CRS_HOME/bin` you may need to change the Oracle user's profile to load `$ORA_CRS_HOME/bin` before `$ORACLE_CRS_HOME/bin` and restart your nodes.

Connection Load Balancing

Using sqlplus or any other program that allows you to log onto the RAC database as SYSDBA, run the following query to display the active sessions and which node in the cluster they are running on:

```
select i.instance_name,
       i.host_name,
       s.username,
       to_char(sysdate, ' Day HH24:MI') current_time,
       s.LAST_CALL_ET seconds_since_active,
       s.failed_over,
       t.sql_text || ' ...' sql_fragment
from   gv$instance i,
       gv$session s,
       gv$sqltext t
where  s.inst_id = i.instance_number
       and s.status = 'ACTIVE'
       and s.username = 'HP_OO_USER'
       and rawtohex(s.sql_address) <> '00'
       and s.sql_address = t.address
       and t.piece = 0;
```

Where **hp_oo_user** is assumed to be the username for the HP OO user. (Note the capitalization in the SQL query.)

For a loaded system, you can see that different instances process queries for the same user.

Useful Information

The Web sites Oracle, Oracle Downloads, and Oracle Base contain a great deal of helpful information, including the following:

- On the Oracle Downloads Web site:
 - Oracle Database High Availability Overview
 - Introduction to Oracle Clusterware and Oracle Real Application Clusters
- On the Oracle Web site:
 - Oracle Application Server 10g Fast Connection Failover Configuration Guide - An Oracle White Paper, October 2006.

This white paper contains a good description of the FCF technology, plus tips for debugging and troubleshooting FAN events. It is written with the Oracle Application Server in mind, but the concepts are still useful in an HP OO environment.

- Workload Management with Oracle Real Application Clusters – An Oracle White Paper.

This white paper includes a good definition of services and descriptions of load balancing advisory, details about the structure of FAN events, useful tips on how to monitor load balancing advisory events.

- The Oracle Base Web site

This site shows how to install RAC on VMware (not for performance benchmarks, but for other testing).

